The Distance Effects on the Intensive and Extensive Margins of Trade Over Time

Juyoung Cheong[*], Do Won Kwak[†], and Kam Ki Tang[‡]

Abstract

The empirical trade literature has long been puzzled by the finding of a large and non-decreasing distance coefficient in the gravity equation amid falling transportation costs over time. To shed new light on this puzzle, the recent theoretical literature shifts its focus to the differential effects of distance on the extensive and intensive margins of trade. However, so far there is a lack of corresponding contributions from studies of empirical gravity equations. This paper provides the first evidence using data for about 150-200 countries between the years 1980 and 2009. Extensive and intensive margins are measured based on bilateral trade data of more than 3100 product items. It is found that the distance effect on the extensive margins declines while that on the intensive margins rises over time. The same conclusion is reached when the distance effects are allowed to be sector-specific.

JEL Code: F14

Keywords: Extensive Margin, Intensive Margin, Distance Effect, Gravity Model

[*]School of Economics, University of Queensland, QLD 4072, Australia; e-mail: j.cheong@uq.edu.au; tel: +61-7-3346-7048; fax: +61-7-3365-7299.

[†]School of Economics, University of Queensland, QLD 4072, Australia; e-mail: d.kwak@uq.edu.au; tel: +61-7-3365-6429; fax: +61-7-3365-7299.

[‡]School of Economics, University of Queensland, QLD 4072, Australia; e-mail: kk.tang@uq.edu.au; tel: +61-7-3365-9796; fax: +61-7-3365-7299.

1 Introduction

A quintessential element of the gravity equation of international trade flows is distance. The assumption of "iceberg" type variable costs in theoretical trade models implies an unambiguous negative distance effect on bilateral trade, which is strongly supported by empirical evidence. With transportation costs per kilometer falling rapidly after the World War II for all ranges of distance but falling faster for longer distance because of advancement in air transport technology (Hummels, 2007), one would expect the effect of geographical distance on international trade flows to reduce over time. However, empirical evidence consistently suggests otherwise: the coefficient of distance in the gravity equation remains the same or even becomes bigger (in absolute terms) over time (Leamer and Levinsohn, 1995; Disdier and Head, 2008).[1] This counter-expectation result comes to be known as the distance puzzle in the trade literature.

To shed new light on this puzzle, the recent theoretical literature shifts its focus to the differential effects of distance on the extensive margin (trade in products not previously traded) and intensive margin (trade in products previously traded) of trade (e.g. Chaney, 2008; Helpman, Melitz, and Rubinstein, 2008; Chaney, 2013). For instance, Krautheim (2012) argues that the distance effect on trade flows can remain large when transportation costs decrease, if the extensive margins become larger with closer countries. The rationale is that, as closer countries trade more for selective products, they accumulate more knowledge of each other's markets; as a result, they face a lower fixed cost in entering each other's untapped product markets, compared with breaking into unfamiliar markets in more distant countries.[2] Bernard, Jensen, Redding, and Schott (2007) find that distance has a strong negative effect on the number of products exported by US firms, but not on the amount they export, supporting the hypothesis that the fixed cost to trade is a positive function of distance. This implies that distinguishing between the extensive and intensive margins of trade and recognizing the differential effects of distance on them could potentially offer a gateway to solving the distance puzzle.

[1] Yotov (2012) provides supporting evidence of globalization by showing that intranational trade gets smaller compared with international trade over time, but his findings do not necessarily imply that long-distance trade costs decrease faster than short distance costs over time.

[2] Also based on the idea of firm networks, Chaney (2013) explains why the distance elasticity of aggregate trade flows could be independent of transportation technology and therefore constant over time.

To our knowledge, so far there have been no empirical studies examining the distance effects on the extensive and intensive margins of trade based on gravity equations. Lin and Sim (2012) provide some evidence of increasing extensive margins at longer distance but increasing intensive margins at shorter distance. However, they do not estimate any gravity equations for the two margins and therefore cannot verify how the distance coefficients on them evolve. The current paper provides the first piece of evidence of such using Standard International Trade Classification (SITC) five-digit bilateral trade data. This is the most dissaggregate data available for the large number of countries and the long time duration considered in this paper.

2 Methodology and Data

2.1 Extensive and Intensive Margins

Following Hummels and Klenow (2005), we measure the extensive margin of exports from county j to county i in year t as

$$EM_{ijt} = \frac{\sum_{p \in \mathbf{P}_{ijt}} T_{iWt}(p)}{\sum_{p \in \mathbf{P}_{iWt}} T_{iWt}(p)} \tag{1}$$

where p is a product index; $\mathbf{P}_{ijt}$ is the set of products that j exports to i; $\mathbf{P}_{iWt}$ is the set of products that all counties in the world (W) export to i; $T_{iWt}(p)$ is the export flows for product p from all counties in the world to i. The numerator of (1) is the number of product varieties that j exports to i, and the denominator is the number of product varieties all countries in the world export to i, with both being weighted by the world export flows of the products to i. In other words, more important products in i's import basket are weighted more heavily. Therefore, EM_{ijt} measures the diversity of j's exports to i compared with the diversity of world's exports to i. By construction, the measure is bounded between 0 and 1, where a larger number means j is present in more of i's import markets.

The intensive margin of exports from j to i in year t is

$$IM_{ijt} = \frac{\sum_{p \in \mathbf{P}_{ijt}} T_{ijt}(p)}{\sum_{p \in \mathbf{P}_{ijt}} T_{iWt}(p)} \tag{2}$$

where $T_{ijt}(p)$ is the export flows for a product p, from j to i. The numerator of (2) is j's total exports to i and the denominator is the world's total exports of the same type of products to i. Therefore, IM_{ijt} measures the overall market share j has for the set of products that it exports to i. The measure is also bounded between 0 and 1, where a larger number means j commands a larger share of i's import markets where it has a presence.

The product of the two margins is

$$EM_{ijt} \cdot IM_{ijt} = \frac{\sum_{p \in \mathbf{P}_{ijt}} T_{ijt}(p)}{\sum_{p \in \mathbf{P}_{iWt}} T_{iWt}(p)} = \frac{T_{ijt}}{\sum_{\forall k} T_{ikt}} \tag{3}$$

where T_{ijt} is aggregate exports from j to i; $EM_{ijt} \cdot IM_{ijt}$ equals the market share of j in i's overall import market.

2.2 Gravity Model

Taking the logs of (3) yields

$$lnT_{ijt} = lnEM_{ijt} + lnIM_{ijt} + lnT_{it.} \tag{4}$$

This indicates that if lnT_{ijt} can be modeled by a gravity equation $g(\mathbf{x})$, i.e. $lnT_{ijt} = g(\mathbf{x})$ where $\mathbf{x}$ is a set of factors, then we can also model its linear components using the same gravity equation up to a scale factor, i.e. $lnEM_{ijt} = ag(\mathbf{x})$, $lnIM_{ijt} = bg(\mathbf{x})$, and $lnT_{jt} = (1-a-b)g(\mathbf{x})$ where $a, b \in \mathbb{R}$. For $g(\mathbf{x})$ we adopt the general equilibrium framework suggested by Anderson and van Wincoop (2003):

$$\begin{aligned} lnEM_{ijt}(or\ lnIM_{ijt}) &= \alpha_0 + \alpha_1 lnY_{it} + \alpha_2 lnY_{jt} + \alpha_3 lny_{it} + \alpha_4 lny_{jt} + \mathbf{X}_{ijt}\beta \\ &\quad + \sum_{t=1}^{T} \gamma_t ln(D_{ij}) \cdot d_t + u_{it} + v_{jt} + \varepsilon_{ijt} \end{aligned} \tag{5}$$

where D_{ij} is the geographical distance between i and j (in km); d_t is year indicator; Y is real GDP measured in purchasing power parity (PPP) terms; y is real per-capita GDP (PPP); $\mathbf{X}$ is a vector containing trade cost variables, dummies for regional trade agreements (RTA), General Agreement on Tariffs and Trade/World Trade Organization (GATT/WTO) memberships, currency union, common border, common language, colony and colonizer, the numbers of island and landlocked countries in a pair (0,1 or 2) and the geographic area of a country; u_{it} and v_{jt} are, respectively, the importer-time and exporter-time fixed effects (FEs) to control for the multilateral resistance terms. The interaction between d and $ln(D_{ij})$ allows the coefficient of the latter to vary over time.

2.3 Data

We use data for every five years, starting from 1980 but finishing at 2009 instead of 2010 due to data constraints. That is, we use data for 1980, 1985, 1990, and so forth until 2005, then 2009. Using data for every five year allows us to control for the multilateral resistance terms using country-time fixed effects. Data on nominal bilateral trade flows are drawn from UNCOMTRADE. Extensive and intensive margins of exporting are constructed based on the latest third revision of SITC five-digit codes, which categorize trade products into 3,121 items. Data on nominal GDP and GDP per capita are drawn from the Penn World Table (PWT) 7.0, and data on GDP deflators are drawn from the U.S. Department of Commerce's Bureau of Economic Analysis. RTA data are constructed from Regional Trade Agreements Information System (RTA-IS) of the WTO and data on GATT/WTO memberships are also drawn from the WTO website. The data on currency unions are an updated and extended version of the list provided by Glick and Rose (2002). Data on distance, common border, common language, colony and colonizer, the numbers of island and landlocked status and the geographic area are from CEPII. Time-varying religion data between 1980 and 2010 are obtained from Maoz and Henderson (2013).

3 Estimation Results

Our first sample covers 71 exporting countries and 158 importing countries and has no zero observations. In this sample, we keep the country-pairs throughout the sample period fixed as in 1980. Carrère and Schiff (2005) and Lin and Sim (2012) suggest that using the rising coefficient for distance in year-by-year cross-country regressions as evidence of increasing impediment of geographical distance on bilateral trade could be misleading. This is because in the early period of a sample, countries farther apart are less likely to trade with each other and therefore are excluded from the estimation. Keeping the country-pair fixed over time can avoid this problem.

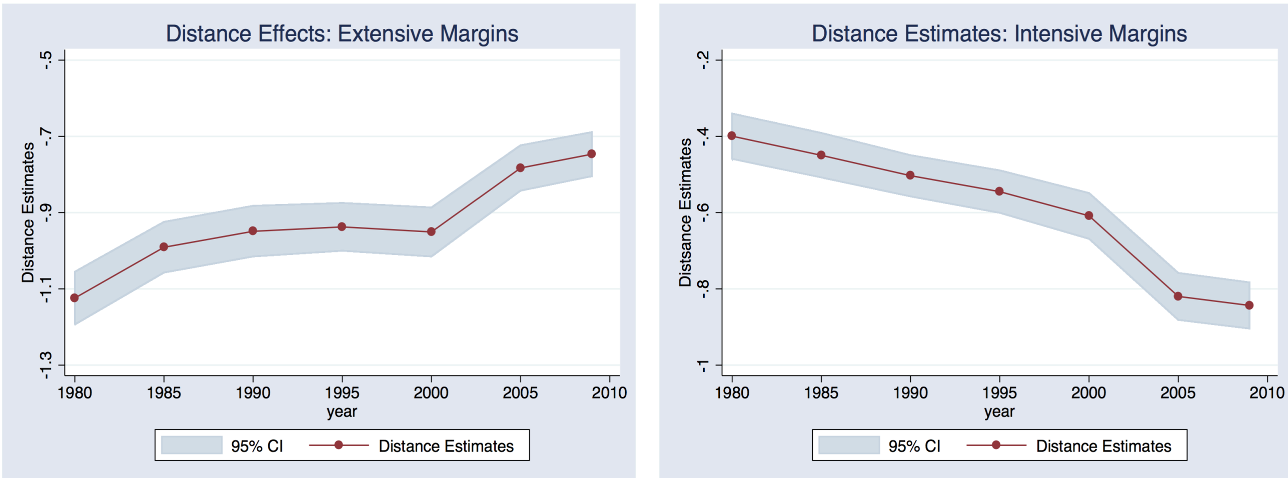
The estimate coefficient on the log distance variable and its 95% confidence intervals are shown in Figure 1. Distance has negative effects on both extensive and intensive margins as expected, meaning that closer countries trade more in variety as well as volume. Until 2005, distance effects are significantly larger for the extensive margins than for the intensive margins.

The distance effect on the extensive margins is unmistakably diminishing (in absolute terms) over time, with the coefficient of log distance reducing by around 15% in total during the three decades. The attenuation of the distance effect is particularly acute since year 2000. Our finding stands opposite to the proposition of Krautheim (2012). One possible explanation for our finding is that the rapid advancement of the Internet and e-commerce, especially in the last decade, has reduced the information and setup costs for exporters, allowing them to enter deeper into previously unfamiliar territories. As such, the impediment of distance on countries' effort to expand their export varieties in those markets has diminished over time.

On the contrary, the distance effect on the intensive margins is increasing (in absolute terms) over time, with the coefficient increasing by an even larger proportion of around 24% during the three decades. Again, the decrease is particularly notable since year 2000. But there are no obvious explanations for the growing distance effect on the intensive margin, echoing the original puzzle of a growing distance effect on bilateral trade flows.

We offer several sensitivity tests for the OLS estimations in Figure 1. First, we estimate the model without restricting the country-pairs to be the same as those with non-zero trade in 1980. This sample

Figure 1: Distance effects: balanced data



(a) Distance and Extensive Margins: OLS

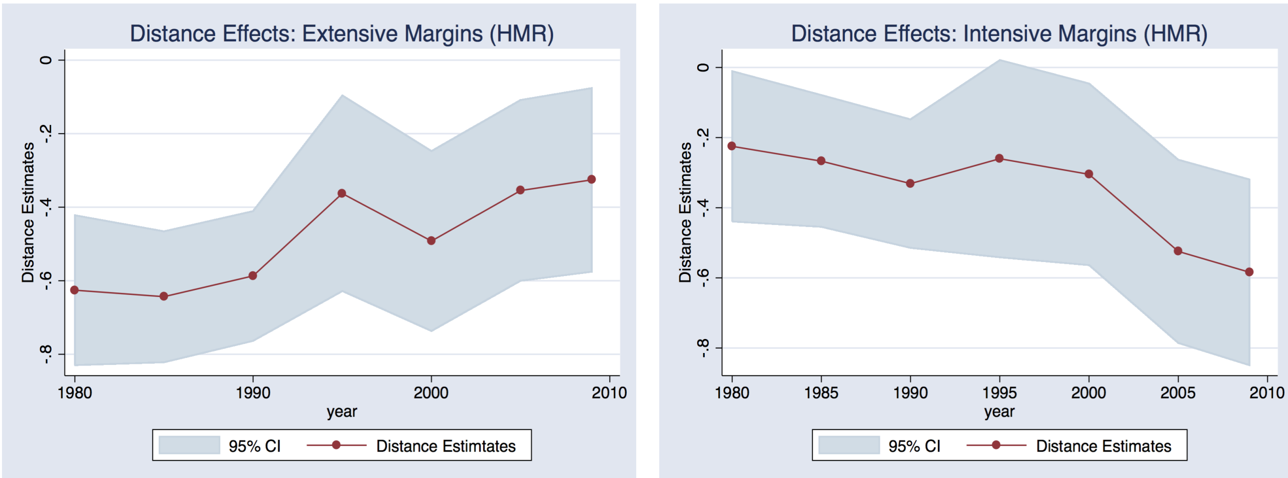
(b) Distance and Intensive Margins: OLS

covers 183 exporting and 195 importing countries. Second, we replace equations (1) and (2) with the conventional measures of extensive and intensive margins. The conventional measure of extensive margins is equal to the number of products j exports to i in year t, and the conventional measure of intensive margins is the average number of exports per product that j exports to i in year t. That is, the conventional measures are absolute, unweighted measures, as in contrast to the relative, weighted measures by Hummels and Klenow (2005). The results of both sensitivity tests are qualitatively the same as Figure 1. To save space, we omit the details here.

Helpman, Melitz, and Rubinstein (2008) (HMR) note that frequently observed "zeros" in the bilateral trade data might be associated with firms' fixed costs of exporting and emphasize the importance of accounting for them to avoid selection bias in the gravity model. In the third sensitivity test, we include zero observations in the sample but keep the country-pairs fixed as in 1980. In this sample, the numbers of exporting and importing countries are 143 and 167 respectively. Following HMR, we use a common religion index as the exclusion restriction variable for their two-stage method.[3] As shown in Figure 2, the point estimates of the distance coefficients are similar to those from the OLS estimates without zero

[3] See Appendix I in Helpman, Melitz, and Rubinstein (2008) for how to construct a common religion index. Our set of religions include the four most popular religions: Christianity, Judaism, Islam and Buddhism. Because religion data are only available every five years, we use the 2010 data for year 2009.

Figure 2: Distance effects: HMR 2-stage estimator



(a) Distance and Extensive Margins: HMR

(b) Distance and Intensive Margins: HMR

observations, although the standard errors are larger.[4]

Krautheim (2012) suggests that the information spillover effect on the fixed costs of international trade can be sector-specific. Therefore, we also estimate distance effects at the sectoral level to see if any sector exhibits properties in line with Krautheim's proposition. Sectors are categorized according to the SITC one-digit codes.[5] The results are shown in Table 1. To save space, we only report the results for every decade. First, distance has significant and negative effects on both extensive and intensive margins for all sectors in all periods. For the extensive margins, except for sectors (3) (crude materials, inedible, except fuels) and (4) (mineral, fuels, lubricants and related materials), all other sectors register a smaller (in absolute terms) distance effect in 2009 than in 1980. For the intensive margins, all sectors register a larger (in absolute terms) distance effect in 2009 than in 1980. In other words, the results from the sectoral analysis are qualitatively consistent with the results in Figure 1.

[4]As the rule of thumb, OLS standard errors and HMR standard errors have a relationship of $\sigma_{OLS} \cdot \frac{1}{|\rho_{I,religion}|} \approx \sigma_{HMR}$, where I is an indicator of positive trade flows, *religion* is the religious proximity variable and ρ is the correlation coefficient. As $|\rho| < 1$ in the sample, we always have $\sigma_{OLS} < \sigma_{HMR}$. Therefore, the weak power of the exclusion restriction measured by ρ leads to large HMR standard errors.

[5]We, however, combine the food and live animals groups and omit the service sector.

Table 1: Distance and Intensive and Extensive Margins: SITC 1 digit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	FOOD AND LIVE ANIMALS	BEVERAGES AND TOBACCO	CRUDE MATERIALS, INEDIBLE, EXCEPT FUELS	MINERAL, FUELS, LUBRICANTS AND RELATED MATERIALS	ANIMAL AND VEGETABLE OILS, FATS AND WAXES	CHEMICALS AND RELATED PRODUCTS	MANUFACTURED GOODS	MACHINERY AND TRANSPORT EQUIPMENT	MISCELLANEOUS MANUFAC-TURED ARTICLES
Extensive Margin									
1980	-0.965***	-0.428***	-0.886***	-0.610***	-1.111***	-0.840***	-1.062***	-0.892***	-0.883***
	(0.033)	(0.037)	(0.046)	(0.051)	(0.088)	(0.039)	(0.037)	(0.035)	(0.034)
1990	-0.819***	-0.495***	-0.932***	-0.827***	-0.825***	-0.915***	-1.037***	-0.811***	-0.721***
	(0.032)	(0.031)	(0.043)	(0.053)	(0.071)	(0.036)	(0.036)	(0.034)	(0.032)
2000	-0.794***	-0.356***	-0.930***	-0.872***	-0.813***	-0.839***	-1.001***	-0.792***	-0.734***
	(0.030)	(0.025)	(0.041)	(0.058)	(0.067)	(0.036)	(0.035)	(0.035)	(0.031)
2009	-0.702***	-0.311***	-0.890***	-0.657***	-0.869***	-0.799***	-0.922***	-0.763***	-0.637***
	(0.028)	(0.025)	(0.042)	(0.044)	(0.067)	(0.036)	(0.033)	(0.033)	(0.029)
Intensive Margin									
1980	-0.492***	-0.650***	-0.482***	-1.055***	-0.179***	-0.769***	-0.623***	-0.639***	-0.614***
	(0.035)	(0.062)	(0.044)	(0.084)	(0.073)	(0.046)	(0.035)	(0.036)	(0.035)
1990	-0.659***	-0.737***	-0.490***	-1.267***	-0.292***	-0.738***	-0.640***	-0.737***	-0.714***
	(0.032)	(0.061)	(0.043)	(0.075)	(0.075)	(0.043)	(0.031)	(0.033)	(0.034)
2000	-0.842***	-1.022***	-0.607***	-1.384***	-0.356***	-0.941***	-0.766***	-0.846***	-0.867***
	(0.035)	(0.056)	(0.044)	(0.094)	(0.074)	(0.052)	(0.034)	(0.037)	(0.038)
2009	-0.964***	-1.247***	-0.778***	-1.563***	-0.514***	-1.048***	-1.077***	-0.996***	-1.059***
	(0.036)	(0.058)	(0.049)	(0.085)	(0.079)	(0.055)	(0.035)	(0.038)	(0.038)

Note: All estimates are obtained accounting for CPFEs and CTFEs. Standard errors in parenthses. ***:significant at 1%

4 Conclusion

We have provided the first evidence of distance effects on the extensive and intensive margins of bilateral trade over time. The two margins are calculated based on data for 3,121 product items, and the data cover 158-195 countries between the years 1980 and 2009. We find that the distance effect on the extensive margins becomes smaller (in absolute terms) over time, while that on the intensive margins becomes larger. The same conclusion is reached even when we further allow the distance effects to vary across sectors.

References

ANDERSON, J. E., AND E. VAN WINCOOP (2003): "Gravity with gravitas: a solution to the border puzzle," *American Economic Review*, 83(1), 170–192.

BERNARD, A. B., J. B. JENSEN, S. J. REDDING, AND P. K. SCHOTT (2007): "Firms in International Trade," *Journal of Economic Perspectives*, 21(3), 105–130.

CARRÈRE, C., AND M. SCHIFF (2005): "On the geography of trade: distance is alive and well," *Revue Economique*, 56, 1249–1274.

CHANEY, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4), 1707–21.

——— (2013): "The Gravity Equation in International Trade: An Explanation," Working Paper 19285, National Bureau of Economic Research.

DISDIER, A.-C., AND K. HEAD (2008): "The Puzzling Persistence of the Distance Effect on Bilateral Trade," *The Review of Economics and Statistics*, 90(1), 37–48.

GLICK, R., AND A. K. ROSE (2002): "Does a currency union affect trade? The time-series evidence," *European Economic Review*, 46(6), 1125–1151.

HELPMAN, E., M. MELITZ, AND Y. RUBINSTEIN (2008): "Estimating Trade Flows: Trading Partners and Trading Volumes," *The Quarterly Journal of Economics*, 123(2), 441–487.

HUMMELS, D. (2007): "Transportation costs and international trade in the second era of globalization," *Journal of Economic Perspective*, 21(3), 131–154.

HUMMELS, D., AND P. J. KLENOW (2005): "The Variety and Quality of a Nation's Exports," *American Economic Review*, 95(3), 704–723.

KRAUTHEIM, S. (2012): "Heterogeneous firms, exporter networks and the effect of distance on international trade," *Journal of International Economics*, 87, 27–35.

LEAMER, E., AND J. LEVINSOHN (1995): *International trade theory: the evidence*pp. 1339–1394, Handbook of International Economics, Volume 3. Elsevier, Amsterdam.

LIN, F., AND N. C. SIM (2012): “Death of distance and the distance puzzle,” *Ecoonomics Letters*, 116, 225–228.

MAOZ, Z., AND E. A. HENDERSON (2013): “The World Religion Dataset, 1945-2010: Logic, Estimates, and Trends,” *International Interactions*, 39(3), 265–291.

YOTOV, Y. V. (2012): “A simple solution to the distance puzzle in international trade,” *Economics Letters*, 117(3), 794–798.